第26卷 第6期　　工 程 数 学 学 报　　Vol. 26 No. 6
2009年12月　　CHINESE JOURNAL OF ENGINEERING MATHEMATICS　　Dec. 2009

文章编号:1005-3085(2009)06-1090-07

对称矩阵反问题的总体最小二乘解*

吕良福[1]， 戴 华[2]， 徐 欢[3]

(1- 天津大学数学系，天津 300072; 2- 南京航空航天大学数学系，南京 210016;
3- 天津市耀华中学数学组，天津 300040)

摘 要: 最小二乘法是近年来求解对称矩阵反问题的一种常用方法，但因系数矩阵常常存在误差，方法本身具有很大的局限性。鉴于此，本文提出并讨论了对称矩阵反问题的总体最小二乘解，给出了解的一般表达式；证明了最佳逼近问题解的存在唯一性，给出了其具体表达式及数值算法，并将数值结果应用于求解对称矩阵反问题。

关键词: 对称矩阵；反问题；总体最小二乘解；奇异值分解；Riccati 方程

分类号: **AMS**(2000) 15A09; 65F10　　**中图分类号**: O241.6　　**文献标识码**: A

1 引言

用 $R^{n\times m}$ 表示 $n\times m$ 实矩阵的全体; $R^n = R^{n\times 1}$; $SR^{n\times n}$ 表示 n 阶实对称矩阵的全体; $OR^{n\times n}$ 表示 n 阶正交矩阵的全体； I_k 表示 k 阶单位阵； A^+ 表示矩阵 A 的 Moore-Penrose 广义逆； $\mathrm{Im}A$ 表示矩阵 A 的值域，即 $\mathrm{Im}A = \{Ax \mid x\in R^n\}$，其中 $A\in R^{n\times n}$； $\sigma(A)$ 表示矩阵 A 的特征值的全体； $\mathrm{tr}(A)$ 表示矩阵 A 的迹； $R(A)$ 表示由 A 的列向量组成的线性组合的全体，即矩阵 A 的值域。$\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 表示对角元素为 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 的 n 阶对角矩阵；在 $R^{n\times m}$ 上定义矩阵 A 与 B 的内积为 $\langle A, B\rangle = \mathrm{tr}(B^TA)$，则由此内积导出的范数 $\|A\| = \sqrt{\langle A, A\rangle}$ 是矩阵的 Frobenius 范数，并且 $R^{n\times m}$ 构成一个完备的内积空间。

文 [1] 讨论了对称矩阵反问题的最小二乘解。文 [2] 从数值分析角度研究了总体最小二乘问题。文 [3,4] 通过分析比较最小二乘与总体最小二乘问题，理论上说明了总体最小二乘是比最小二乘更合适的拟合方法。

本文研究如下问题:

问题 I 给定 $X, B\in R^{n\times m}$，求 $A\in SR^{n\times n}$，使得 $A\hat{X} = \hat{B}$，其中 $\hat{X}, \hat{B}$ 满足

$$\min \left\|[X-\hat{X},\, B-\hat{B}]\right\|, \quad \text{s.t.} \quad \mathrm{rank}(\hat{X}) = \mathrm{rank}\begin{pmatrix} \hat{X} \\ \hat{B} \end{pmatrix}.$$

且 $SR^{n\times n} = \{A \mid A\in R^{n\times n},\, A^T = A\}$。

问题 II 给定 $\tilde{A}\in R^{n\times n}$，求 $\hat{A}\in S_A$，使得

$$\|\tilde{A}-\hat{A}\| = \inf_{A\in S_A} \|\tilde{A}-A\|, \tag{1}$$

其中 S_A 是问题 I 的解集合。

收稿日期: 2007-10-30. **作者简介**: 吕良福 (1979年7月生)，男，博士，讲师. 研究方向：计算机图形学、数值代数.
***基金项目**: 国家自然科学基金 (60673196)；国家自然科学青年基金 (60704015).

本文首先给出了问题I有解的充分条件，给出了解的一般表达式；其次证明了问题II最佳逼近解的存在唯一性，给出了其表达式及求解算法；最后将数值结果应用于求解对称矩阵反问题。

2 问题I的解

首先引入几个概念。

定义2.1[5] 设M是R^n的一个子空间，$M \subseteq R^n$，称M是H-中性的，如果

$$[x, y] = \langle Hx, y\rangle = y^T Hx = 0, \quad \forall\, x, y \in M.$$

定义2.2[5] 对于$m \times n$实矩阵X、n阶单位阵I_n，记

$$G(X) = \mathrm{Im}\begin{pmatrix} I_n \\ X \end{pmatrix} \subseteq R^{m+n},$$

我们把n维子空间$G(X)$称作X的图。进一步，对于R^{m+n}中的n维子空间S，若S是某个$m \times n$矩阵的图，则称S为图子空间。

下面考虑矩阵方程$CY \approx D(C, D \in R^{m\times n})$，它的带对称约束的总体最小二乘问题是求$Y \in SR^{n\times n}$，使$\hat{C}Y = \hat{D}$，其中$\hat{C}, \hat{D}$满足

$$\min \left\|[C - \hat{C},\, D - \hat{D}]\right\|, \quad \text{s.t.} \quad R(\hat{D}) \subset R(\hat{C}).$$

利用Lagrange乘子法，文[5]证明了矩阵方程$CY \approx D$带对称约束的总体最小二乘问题可归结为求Riccati矩阵方程$XSX + RX + XR - S = 0$的解，其中$R = C^TC - D^TD$, $S = C^TD - D^TC$。该方程的解与矩阵

$$T = \begin{pmatrix} R & S \\ S & -R \end{pmatrix}$$

的不变子空间密切相关。记

$$K = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

T是实对称矩阵，且$KT = -TK$，则存在非奇异实矩阵W使得

$$W^{-1}TW = \mathrm{diag}[\lambda_k,\, -\lambda_k,\, \cdots,\, \lambda_1,\, -\lambda_1,\, 0,\, \cdots,\, 0],$$

其中$\lambda_k \geq \lambda_{k-1} \geq \cdots \geq \lambda_1 > 0$是$T$的正特征值[5]。

引理2.1[5] 设$C, D \in R^{m\times n}$，令

$$T = \begin{pmatrix} R & S \\ S & -R \end{pmatrix},$$

其中$R = C^TC - D^TD$, $S = C^TD + D^TC$，且非奇异实矩阵W使得

$$W^{-1}TW = \mathrm{diag}[\lambda_k,\, -\lambda_k,\, \cdots,\, \lambda_1,\, -\lambda_1,\, 0,\, \cdots,\, 0],$$

其中$\lambda_k \geq \lambda_{k-1} \geq \cdots \geq \lambda_1 > 0$是$T$的正特征值，则存在$n$维的、$T$不变的、$K$-中性的子空间$S_0$，使得$\sigma(T \mid S_0) = \{\lambda_k, \cdots, \lambda_1, 0\}$；或如果$k = n$，有$\sigma(T \mid S_0) = \{\lambda_k, \cdots, \lambda_1\}$。

引理2.2[5] 若引理2.1中S_0是图子空间，则由它可确定关于$[C, D]$的带对称约束的总体最小二乘解，特别地，当$k = n$时，带对称约束的总体最小二乘解是唯一的。

引理2.3[5] 带对称约束的总体最小二乘问题中修正矩阵$\hat{C}$满足$\hat{C} = (C+DY)(I+Y^2)^{-1}$，其中$Y$是对称总体最小二乘解。

利用带对称约束总体最小二乘问题的结果，可得对称矩阵反问题的总体最小二乘解。

定理2.1 设$X, B \in R^{n\times m}$，令

$$T = \begin{pmatrix} R & S \\ S & -R \end{pmatrix}, \quad R = X^TX - B^TB, \quad S = X^TB + B^TX, \tag{2}$$

且存在非奇异实矩阵W使得

$$W^{-1}TW = \mathrm{diag}[\lambda_k, -\lambda_k, \cdots, \lambda_1, -\lambda_1, 0, \cdots, 0],$$

其中$\lambda_k \geq \lambda_{k-1} \geq \cdots \geq \lambda_1 > 0$是$T$的正特征值，由$W$的奇数列组成的子阵$\begin{pmatrix} Y \\ Z \end{pmatrix}$，若$Y \in R^{n\times n}$非奇异，则$A = ZY^{-1}$为问题I的一个解，且问题I解的一般表达式为

$$A = A_0 + (I - \hat{X}\hat{X}^+)M(I - \hat{X}\hat{X}^+), \quad \forall\, M \in SR^{n\times n}, \tag{3}$$

其中$A_0 = ZY^{-1}$，$\hat{X} = (I + {A_0}^2)^{-1}(X + A_0B)$。

证明 设问题I的任一解$A \in SR^{n\times n}$，则$\hat{X}^TA = \hat{B}^T$，其中$\hat{X}, \hat{B}$满足

$$\min \left\|[X^T - \hat{X}^T, B^T - \hat{B}^T]\right\|, \quad \text{s.t.} \quad R(\hat{B}^T) \subset R(\hat{X}^T),$$

即A可看作关于$[X^T, B^T]$的对称总体最小二乘解。

在引理2.1中令$C = X^T$，$D = B^T$，得

$$R = X^TX - B^TB, \quad S = X^TB + B^TX.$$

由引理2.1知，矩阵$\begin{pmatrix} Y \\ Z \end{pmatrix}$的列张成的子空间即为满足引理2.1条件的子空间$S_0$。

又因S_0是图子空间的充要条件是Y^{-1}存在，所以由引理2.2知$A = ZY^{-1}$为问题I的解。

结合引理2.3，下面只需证明$\bar{A} = (I - \hat{X}\hat{X}^+)M(I - \hat{X}\hat{X}^+)$是方程$A\hat{X} = 0\,(A \in SR^{n\times n})$的通解。

事实上，对任意的$M \in SR^{n\times n}$，$\bar{A} = (I - \hat{X}\hat{X}^+)M(I - \hat{X}\hat{X}^+)$是方程$A\hat{X} = 0\,(A \in SR^{n\times n})$的解；而对$A\hat{X} = 0$的任意一个解$\bar{A} \in SR^{n\times n}$，令$M = \bar{A}$，有

$$(I - \hat{X}\hat{X}^+)\bar{A}(I - \hat{X}\hat{X}^+) = (I - \hat{X}\hat{X}^+)\bar{A} = \bar{A} - \hat{X}\hat{X}^+\bar{A} = \bar{A} - (\hat{X}^+)^T(\hat{X})^T\bar{A} = \bar{A},$$

所以$\bar{A} = (I - \hat{X}\hat{X}^+)M(I - \hat{X}\hat{X}^+)$是方程$A\hat{X} = 0\,(A \in SR^{n\times n})$的通解。

于是，问题I解的一般表达式为

$$A = A_0 + (I - \hat{X}\hat{X}^+)M(I - \hat{X}\hat{X}^+), \quad \forall\, M \in SR^{n\times n},$$

其中$A_0 = ZY^{-1}$，$\hat{X} = (I + {A_0}^2)^{-1}(X + A_0B)$。

3　问题 II 的解

引理 3.1[6]　设 $E \in R^{n\times n}$，则对任意 n 阶实对称矩阵 H 有

$$\left\|E - \frac{E+E^T}{2}\right\| \le \|E - H\|.$$

定理 3.1　对给定的 $\tilde{A} \in R^{n\times n}$，以及满足定理 2.1 条件的 $X, B \in R^{n\times m}$，问题 II 存在唯一解 $\hat{A} \in S_A$。如果 $\hat{X}$ 的奇异值分解为

$$\hat{X} = U\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}V^T,$$

其中

$$U = (U_1\ \ U_2) \in OR^{n\times n},\quad V = (V_1\ \ V_2) \in OR^{m\times m},$$
$$U_1 \in R^{n\times r},\quad V_1 \in R^{m\times r},\quad r = \mathrm{rank}(\hat{X}),$$

记

$$U^T(\tilde{A} - A_0)U = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix},\quad \tilde{A_{11}} \in R^{r\times r},$$

则问题 II 的解 $\hat{A}$ 可表示为

$$\hat{A} = A_0 + \frac{1}{2}U_2(\tilde{A}_{22} + \tilde{A}_{22}^T){U_2}^T. \tag{4}$$

证明　显然

$$S_A = \left\{A_0 + (I - \hat{X}\hat{X}^+)M(I - \hat{X}\hat{X}^+) \mid A_0 = ZY^{-1},\ \forall\, M \in SR^{n\times n}\right\}$$

是完备内积空间 $R^{n\times n}$ 中的一个闭凸集，由最佳逼近定理[7] 知存在唯一的 $\hat{A} \in S_A$，使式 (1) 成立。

下面证明问题 II 的解 $\hat{A}$ 可由 (4) 式表示。对 $A \in S_A$，记

$$U^TMU = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix},\quad M_{11} \in R^{r\times r},$$

则

$$\begin{aligned}
\|\tilde{A} - A\|^2 &= \left\|\tilde{A} - A_0 - (U_2{U_2}^T)M(U_2{U_2}^T)\right\|^2 \\
&= \left\|U^T(\tilde{A} - A_0)U - \begin{pmatrix} 0 & 0 \\ 0 & I_{n-r} \end{pmatrix}U^TMU\begin{pmatrix} 0 & 0 \\ 0 & I_{n-r} \end{pmatrix}\right\|^2 \\
&= \left\|U^T(\tilde{A} - A_0)U - \begin{pmatrix} 0 & 0 \\ 0 & M_{22} \end{pmatrix}\right\|^2 \\
&= \|\tilde{A}_{11}\|^2 + \|\tilde{A}_{12}\|^2 + \|\tilde{A}_{21}\|^2 + \|\tilde{A}_{22} - M_{22}\|^2.
\end{aligned}$$

由引理3.1可知，当

$$M_{22}=\frac{\tilde{A}_{22}+\tilde{A}_{22}^T}{2} \tag{5}$$

时，$\|\tilde{A}-A\|$ 达到最小值，由(3)和(5)式，可得

$$A=A_0+\frac{1}{2}U_2(\tilde{A}_{22}+\tilde{A}_{22}^T){U_2}^T.$$

算法3.1 求问题II的解的步骤如下：

步1: 按(2)式计算T；

步2: 由定理2.1计算问题I的一个解 $A_0=ZY^{-1}$，即计算 $\begin{pmatrix} Y \\ Z \end{pmatrix}$，并判断 Y^{-1} 是否存在；

步3: 求 $\hat{X}$，并对其进行奇异值分解；

步4: 利用(4)式求 $\hat{A}$。

4 数值算例

下面通过例子说明前述理论和算法的应用。

例 设 $n=4$, $m=3$，给定如下的 $X, B\in R^{n\times m}$，$\tilde{A}\in R^{n\times n}$，求问题II的解。

$$X=\begin{pmatrix} 2.5 & 3 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 0.5 \\ 1.5 & 2 & 0 \end{pmatrix},\quad B=\begin{pmatrix} 3 & 1.5 & 0 \\ 0 & 0 & 1.5 \\ -1 & 0 & 0 \\ 3.5 & 1 & 0 \end{pmatrix},\quad \tilde{A}=\begin{pmatrix} 3 & -2 & 0 & 2 \\ 2 & 1 & 0 & -1 \\ 1 & 2 & 1 & -3 \\ -1 & 1 & 2 & 2 \end{pmatrix}.$$

计算

$$T=\begin{pmatrix} R & S \\ S & -R \end{pmatrix},$$

其中

$$R=\begin{pmatrix} 4 & 3 & 0.5 & -2.25 \\ 3 & -0.25 & 0.5 & 2 \\ 0.5 & 0.5 & 0.25 & 2 \\ -2.25 & 2 & 2 & -7 \end{pmatrix},\quad S=\begin{pmatrix} 24 & 1.5 & -5.5 & 19.25 \\ 1.5 & 3 & 0.75 & 1 \\ -5.5 & 0.75 & 2 & -5 \\ 19.25 & 1 & -5 & 14.5 \end{pmatrix}.$$

计算 $\begin{pmatrix} Y \\ Z \end{pmatrix}$，其中

$$Y=\begin{pmatrix} 0.1038 & 0.1724 & 0.1688 & 0.5806 \\ 0.1205 & -0.0552 & 0.5822 & 0.0812 \\ 0.1839 & 0.3568 & 0.4284 & -0.1108 \\ -0.2323 & 0.0087 & 0.0104 & 0.3340 \end{pmatrix},$$

$$Z=\begin{pmatrix} 0.3573 & 0.4740 & 0.0486 & 0.4891 \\ -0.1417 & -0.4588 & 0.6369 & -0.0257 \\ 0.6450 & 0.4457 & 0.0772 & -0.1496 \\ 0.5685 & -0.4539 & -0.1870 & 0.5198 \end{pmatrix}.$$

易知 Y^{-1} 存在，于是计算

$$A_0=\begin{pmatrix} 0.0033 & -0.8021 & 1.1531 & 2.0362 \\ -0.8021 & 1.7844 & -0.6386 & 0.6716 \\ 1.1531 & -0.6386 & 0.6442 & -2.0827 \\ 2.0362 & 0.6716 & -2.0827 & -2.8378 \end{pmatrix}.$$

计算 $\hat{X}$，并对其进行奇异值分解

$$\hat{X}=\begin{pmatrix} 3.0053 & 2.4692 & 0.1909 \\ 0.1400 & 0.8430 & 1.0500 \\ -1.0436 & 0.2744 & 0.5502 \\ 1.8263 & 1.2674 & 0.0091 \end{pmatrix}$$

$$=\begin{pmatrix} -0.8510 & 0.0347 & -0.2555 & -0.4576 \\ -0.1496 & 0.7466 & 0.6476 & -0.0269 \\ 0.1365 & 0.6557 & -0.7163 & 0.1958 \\ -0.4846 & -0.1066 & 0.0470 & 0.8669 \end{pmatrix}$$

$$\begin{pmatrix} 4.5746 & 0 & 0 \\ 0 & 1.5330 & 0 \\ 0 & 0 & 0.3607 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} -0.7882 & -0.4372 & 0.4330 \\ -0.6130 & 0.4957 & -0.6153 \\ -0.0544 & 0.7504 & 0.6587 \end{pmatrix}$$

利用 (4) 式计算 $\hat{A}$ 得

$$\hat{A}=\begin{pmatrix} 1.3729 & -0.7218 & 0.5671 & -0.5585 \\ -0.7218 & 1.7892 & -0.6731 & 0.5191 \\ 0.5671 & -0.6731 & 0.8950 & -0.9725 \\ -0.5585 & 0.5191 & -0.9725 & 2.0777 \end{pmatrix}.$$

通过求解得：$\|\hat{A}_{\mathrm{TLS}}-\tilde{A}\|=6.5317$，$\|\hat{A}_{\mathrm{LS}}-\tilde{A}\|=10.8546$，其中 $\hat{A}_{\mathrm{TLS}}$、$\hat{A}_{\mathrm{LS}}$ 为分别运用总体最小二乘法和最小二乘法求解的最佳逼近解。

参考文献：

[1] 戴华. 用振动试验最优校正刚度、柔度和质量矩阵[J]. 振动工程学报，1988, 1(2): 18-27

[2] Golub G H, Van Loan C F. An analysis of the total least squares problem[J]. SIAM Journal of Numerical Analysis, 1980, 17(6): 883-893

[3] Van Huffel S, Vandewalle J. Algebraic connections between the least squares and total least squares problems[J]. Numerical Mathematics, 1989, 55(4): 431-449

[4] Wei M S. Algebraic relations between the total least squares and least squares problems with more than one solution[J]. Numerical Mathematics, 1992, 62(1): 123-148

[5] Lancaster P, Rodman L. Algebraic Riccati Equations[M]. Oxford: Oxford University Press, 1995

[6] Fan K, Hoffman A J. Some metric inequalities in the space of matrices[C]// Proceedings of the American Mathematical Society, 1955, 6: 111-116

[7] 戴华. 矩阵论[M]. 北京：科学出版社，2001

[8] 周树荃，戴华. 代数特征值反问题[M]. 郑州：河南科学技术出版社，1991

The Total Least Squares Solution for the Inverse Problems of Symmetric Matrices

LV Liang-fu[1], DAI Hua[2], XU Huan[3]

(1- Department of Mathematics, Tianjin University, Tianjin 300072;
2- Department of Mathematics, Nanjing University of Aeronautics & Astronautics, Nanjing 210016;
3- Yaohua High School, Tianjin 300040)

Abstract: Least squares have been widely used in the inverse problems of symmetric matrices. However, errors always occur in the relevant coefficient matrix. This makes the approach limited. In order to overcome this shortcoming, the total least squares solution of inverse problems of symmetric matrices are proposed. The general form of the solution is given. The existence and expression of the optimal approximation solution are presented, and a numerical algorithm is derived. These results are finally applied to solve the inverse problem of symmetric matrices.

Keywords: symmetric matrix; inverse problem; total least squares solution; singular value decomposition; Riccati equation